J Fourier Anal Appl (2008) 14: 609–628
DOI 10.1007/s00041-008-9034-0

Weak Convergence of Greedy Algorithms in Banach Spaces

S.J. Dilworth · Denka Kutzarova · Karen L. Shuman · V.N. Temlyakov · P. Wojtaszczyk

Received: 30 March 2007 / Published online: 16 September 2008

Abstract We study the convergence of certain greedy algorithms in Banach spaces. We introduce the WN property for Banach spaces and prove that the algorithms converge in the weak topology for general dictionaries in uniformly smooth Banach

Communicated by Ronald A. DeVore.

The research of the first author was supported by NSF Grant DMS 0701552. The research of the third author was supported by the Grinnell College Committee for the Support of Faculty Scholarship. The research of the fourth author was supported by NSF Grant DMS 0554832 and by ONR Grant N00014-91-J1343. The research of the fifth author was supported by KBN grant 1 PO03A 038 27 and the Foundation for the Polish Science. The fifth author thanks the Industrial Mathematics Institute of the University of South Carolina and Prof. R. DeVore and Prof. V.N. Temlyakov personally for their hospitality. The first and second authors were supported by the Concentration Week on Frames, Banach Spaces, and Signals Processing and by the Workshop in Analysis and Probability at Texas A&M University in 2006.

Current address of D. Kutzarova: Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

S.J. Dilworth (✉) · V.N. Temlyakov
Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA
e-mail: dilworth@math.sc.edu

V.N. Temlyakov
e-mail: temlyak@math.sc.edu

D. Kutzarova
Institute of Mathematics, Bulgarian Academy of Sciences, Sofia, Bulgaria
e-mail: denka@math.uiuc.edu

K.L. Shuman
Department of Mathematics, Grinnell College, Iowa 50112, USA
e-mail: shumank@math.grinnell.edu

P. Wojtaszczyk
Institute of Applied Mathematics, Warsaw University, Banacha 2, 02-097 Warsaw, Poland
e-mail: wojtaszczyk@mimuw.edu.pl

spaces with the WN property. We show that reflexive spaces with the uniform Opial property have the WN property. We show that our results do not extend to algorithms which employ a 'dictionary dual' greedy step.

Keywords Greedy algorithms · Weak convergence · Uniformly smooth Banach spaces · Uniform Opial property

Mathematics Subject Classification (2000) Primary 41A65 · Secondary 42A10 · 46B20

1 Introduction

We study convergence in the weak topology of different greedy algorithms acting in a uniformly smooth Banach space. The first result on convergence of a greedy algorithm is the result of Huber [10]. He proved convergence of the Pure Greedy Algorithm in the weak topology of a Hilbert space H and conjectured that the Pure Greedy Algorithm converges in the strong sense (in the norm of H). L. Jones [12] proved this conjecture. The theory of greedy approximation in Hilbert spaces is now well developed (see [20]). Our interest in this paper is in convergence results for greedy approximation in Banach spaces. There is a number of open problems on convergence (in the strong sense) of different greedy algorithms in Banach spaces (see [20]).

Let X be a Banach space. We say that $\mathcal{D} \subset X$ is a *dictionary* if the following conditions are satisfied: (i) the linear span of $\mathcal{D}$ is norm-dense in X; (ii) $\|g\| = 1$ for all $g \in \mathcal{D}$; (iii) $\mathcal{D}$ is symmetric, i.e. $g \in \mathcal{D} \iff -g \in \mathcal{D}$.

For some algorithms (the Weak Chebyshev Greedy Algorithm ([6, 19]), the Weak Greedy Algorithm with Free Relaxation ([22]) it is known that the uniform smoothness of X guarantees strong convergence of these algorithms for each element $f \in X$ and any dictionary $\mathcal{D}$ in X. For other algorithms (the Weak Dual Greedy Algorithm, the X-Greedy Algorithm) it is not known if uniform smoothness (even uniform smoothness with a power type modulus of smoothness) guarantees convergence in the strong sense for each $f \in X$ and all $\mathcal{D}$.

There is a result of M. Ganichev and N. Kalton [8] that establishes strong convergence of the Weak Dual Greedy Algorithm in a uniformly smooth Banach space satisfying an extra condition Γ. In this paper we impose the following extra condition (the WN property) on a uniformly smooth Banach space and prove the weak convergence of some greedy algorithms. This condition was used implicitly in [6].

For $f \in X$, $f \neq 0$, let $F_f \in X^*$ denote a norming functional of f, i.e, $\|F_f\|_{X^*} = 1$ and $F_f(f) = \|f\|_X$.

Definition 1.1 X has the WN property if every sequence $\{x_n\} \subset X$, with $\|x_n\|_X = 1$, is weakly null in X whenever the sequence $\{F_{x_n}\}$ is weakly null in X^*.

We prove weak convergence of the X-Greedy Algorithm for a uniformly smooth Banach space satisfying the WN property. We note that there are no results on strong convergence of the X-Greedy Algorithm in nontrivial Banach spaces (infinite-dimensional nonhilbertian spaces). We give an example that demonstrates that the WN property does not imply the Γ property.

The use of the norming functional F_{f_m} of a residual f_m of a greedy algorithm after m iterations for a search of an element from the dictionary to be added in approximation has proved to be very natural. Norming functionals for residuals are used in dual type greedy algorithms. At a greedy step of a dual type greedy algorithm we look for an element $\varphi_m \in \mathcal{D}$ satisfying

$$F_{f_{m-1}}(\varphi_m) \geq t \sup_{g\in\mathcal{D}} F_{f_{m-1}}(g), \quad t \in (0,1]. \tag{1}$$

In this paper we discuss the following modification of the greedy step (1): we look for an element $\varphi_m \in \mathcal{D}$ satisfying

$$F_{\varphi_m}(f_{m-1}) \geq t \sup_{g\in\mathcal{D}} F_g(f_{m-1}). \tag{2}$$

It is known (see [21]) that the modification (2) is useful in the problem of exact recovery of sparsely represented elements in the case of dictionaries with small coherence. We show here that the modification (2) is not good for convergence of greedy approximations with regard to general dictionaries.

Let us now describe the organization of the paper. Section 2 recalls the relevant definitions from Banach space theory. In Sect. 3 we recall the definitions of the greedy algorithms which are considered here and we present our main results on the weak convergence of these algorithms in uniformly smooth spaces with the WN property. We also show how the algorithms can easily be modified to obtain strong convergence. In Sect. 4 we show that many classes of Banach spaces have the WN property, e.g. the class of reflexive spaces with the uniform Opial property, and we exhibit examples of spaces which have the WN property but not the property Γ introduced by Ganichev and Kalton [8]. Finally, we show that our results for general dictionaries do not apply to algorithms which use (2) as the greedy step. In fact, we show that if X is not isometric to a Hilbert space (and satisfies additional mild regularity conditions) then there exists a dictionary in X for which such algorithms break down quite badly.

2 Definitions and Notation

We use standard Banach space notation and terminology as in [15]. In this section we record some terminology which may not be completely standard.

2.1 Bases of Banach Spaces and Related Notions

Let $\{e_i\}$ be a (Schauder) basis for a real Banach space X and let $\{e_i^*\}$ be its sequence of biorthogonal functionals. For $x, y \in X$, we write $x < y$ if the support of x is "to the left" of the support of y, i.e. if

$$\max\{n \in \mathbb{N}\colon e_n^*(x) \neq 0\} < \min\{n \in \mathbb{N}\colon e_n^*(y) \neq 0\}.$$

Definition 2.1 A basis $\{e_i\}$ for a Banach space X is *uniformly reverse monotone* (*URM*) if there exists a function $\theta\colon (0,\infty)\to(0,\infty)$ such that for all $x<y$, with $\|y\|\le 1$ and $\|x\|\ge\varepsilon$, we have

$$\|y\|\le\|x+y\|-\theta(\varepsilon).$$

In Definition 2.1, the function θ may and shall be taken to be nondecreasing.

Let $P_m : X\to \mathrm{Span}\{e_i\}_{i=1}^m$ denote the standard projection onto the span of the first m basis elements and let $Q_m = I - P_m$. The basis is *monotone* if $\|P_n\|=1$ for all $n\ge 1$. Note that if $y_n \stackrel{w}{\to} 0$, then for fixed $m\in\mathbb{N}$,

$$\lim_{n\to\infty}\|P_m(y_n)\|=0. \tag{3}$$

In much of what follows, we shall consider 1-*unconditional bases*; these bases are unconditional bases with unconditional basis constant equal to 1, i.e. such that $\|\sum a_ie_i\| = \|\sum \pm a_ie_i\|$ for all choices of scalars and signs. We note that when $\{e_i\}$ is a 1-unconditional basis, then $\|\sum_{i\in A}a_ie_i\|\le\|\sum a_ie_i\|$ for all $A\subset\mathbb{N}$ and scalars $\{a_i\}$.

2.2 Smoothness Properties of Banach Spaces

Here we recall some standard definitions pertaining to smoothness. Suppose X is a real Banach space. Define for $f,g\in X$ and for $u\in\mathbb{R}$

$$\rho_{f,g}(u)=\frac{\|f+ug\|+\|f-ug\|-2\|f\|}{2}.$$

The point $f\in X$ is a *point of Gâteaux smoothness* if

$$\lim_{u\downarrow 0}\frac{\rho_{f,g}(u)}{u}=0\quad\text{ for all } g\in X.$$

By the Hahn-Banach theorem, each $f\in X$ has at least one norming functional F_f. It is known that f is a point of Gâteaux smoothness if and only if its associated norming functional F_f is unique. We say that a Banach space X is *smooth* (or X *has a Gâteaux differentiable norm*) when every nonzero point in X is a point of Gâteaux smoothness. See [5] for more information on smoothness. The assumption that every nonzero point of X is a point of Gâteaux smoothness guarantees that for *any* nonzero $f\in X$ and *any* dictionary $\mathcal{D}$,

$$\inf_{\substack{t\in\mathbb{R}\\ g\in\mathcal{D}}}\|f-tg\|<\|f\|.$$

Indeed, simply choose $g\in\mathcal{D}$ such that $F_f(g)>0$. Then

$$\|f-tg\|=\|f\|-tF_f(g)+o(t),$$

so $\|f - tg\| < \|f\|$ for all sufficiently small $t > 0$. We will also consider uniformly smooth Banach spaces. For each $\tau > 0$, the *modulus of smoothness* of a Banach space X is defined by

$$\rho_X(\tau) = \sup\left\{\frac{\|x+y\| + \|x-y\|}{2} - 1 : x, y \in X, \|x\| = 1, \|y\| = \tau\right\}$$

[16, p. 59]. We say that X is *uniformly smooth* if $\lim_{\tau\to 0} \frac{\rho_X(\tau)}{\tau} = 0$.

There are Gâteaux smooth Banach spaces which are not uniformly smooth. It is known that a uniformly smooth Banach space is *uniformly Fréchet differentiable*, i.e.

$$\|x + y\| = 1 + F_x(y) + \varepsilon(x, y)\|y\| \tag{4}$$

for all $x, y \in X$ with $\|x\| = 1$, where $\varepsilon(x, y) \to 0$ uniformly for $(x, y) \in \{x : \|x\| = 1\} \times X$ as $\|y\| \to 0$.

2.3 Uniform Opial Property

This is a property which relates weak convergence with the geometry of the unit sphere $S_X := \{x \in X : \|x\| = 1\}$. We say that a Banach space has the *uniform Opial property* if there exists a function $\tau(\varepsilon) > 0$ defined for all $\varepsilon > 0$ such that for all sequences $y_n \stackrel{w}{\to} 0$, with $\|y_n\| = 1$, and for all $x \neq 0$ we have

$$\liminf_{n\to\infty} \|x + y_n\| \geq 1 + \tau(\|x\|). \tag{5}$$

Clearly we may and shall assume that the function τ is nondecreasing. Also we can obviously replace the condition $\|y_n\| = 1$ in the definition by the condition $\lim_{n\to\infty} \|y_n\| = 1$. This property was introduced in [17] and turned out to be useful in geometric fixed point theory (see e.g. [18]).

3 Weak Convergence of Greedy Algorithms

3.1 The Weak Dual Greedy Algorithm (WDGA)

We first review the WDGA from [20, p. 66] (see also [6, p. 491]). Let $\mathcal{D}$ be a dictionary for X, and let $0 < t \leq 1$. To define the WDGA, first define $f_0^D := f_0^{D,t} := f$. Then, for each $m \geq 1$, inductively define

(1) $\phi_m^D := \phi_m^{D,t} \in \mathcal{D}$ to be any element of $\mathcal{D}$ satisfying

$$F_{f_{m-1}^D}(\phi_m^D) \geq t \sup_{g\in\mathcal{D}} F_{f_{m-1}^D}(g).$$

(2) Define a_m via

$$\|f_{m-1}^D - a_m\phi_m^D\| = \min_{a\in\mathbb{R}} \|f_{m-1}^D - a\phi_m^D\|.$$

(3) Denote

$$f_m^D := f_m^{D,t} := f_{m-1}^D - a_m \phi_m^D.$$

Using the same notation as in [6, p. 490], we define G_n^D to be

$$G_n^D = f - f_n^D$$

and

$$f_0 - G_n^D = f - G_n^D = f_n^D.$$

The sequence $\{\| f_m^D \|\}$ is nonnegative and nonincreasing since

$$\min_{a\in\mathbb{R}} \| f_m^D - a\phi_{m+1}^D \| \le \| f_m^D - 0\phi_{m+1}^D \|.$$

Therefore, $\{\| f_m \|\}$ converges to some $\alpha \ge 0$. We say that the algorithm converges strongly if $\alpha = 0$. We note that the sequence $\{G_n^D\}$ resulting from the WDGA will be bounded above in norm:

$$\|G_n^D\| - \|f\| \le \|G_n^D - f\| = \|f_n\| \le \|f\|.$$

In [6], Dilworth, Kutzarova, and Temlyakov prove that the WDGA converges strongly when applied to dictionaries $\{\pm\phi_n\}$, where $\{\phi_n\}$ is a strictly suppression 1-unconditional basis for a Banach space X which has a Fréchet-differentiable norm. Also it follows from [6, Theorem 6] that if X is uniformly smooth and has the WN property then the WDGA converges weakly for any dictionary $\mathcal{D}$. The strong convergence of the WDGA was proved under a different hypothesis in [8].

3.2 The X-Greedy Algorithm (XGA) and the Weak X-Greedy Algorithm (WXGA)

Let $0 < t \le 1$ be a weakness parameter. The following algorithm is called the WXGA when $0 < t < 1$ and is called the XGA when $t = 1$. In the case of the XGA we have to add the assumption that the infimum in the first step is attained.

Let $\mathcal{D}$ be a dictionary for X.

(1) Choose $\phi_m \in \mathcal{D}$ and $\lambda_m \in \mathbb{R}$ such that

$$\|f_{m-1}\| - \|f_{m-1} - \lambda_m\phi_m\| \ge t\Big(\|f_{m-1}\| - \inf_{\substack{\lambda\in\mathbb{R}\\ \phi\in\mathcal{D}}} \|f_{m-1} - \lambda\phi\|\Big).$$

(2) Define $f_m = f_{m-1} - \lambda_m\phi_m$. We call f_m the mth residual of f. The sequence $\{\|f_n\|\}$ is decreasing.

(3) Set $G_m := f - f_m$.

Theorem 3.1 *Suppose that X is uniformly smooth and has the WN property. When the XGA or the WXGA is applied to $f \in X$, then $f_n \stackrel{w}{\to} 0$.*

Proof If $\|f_n\| \downarrow 0$ then f_n converges strongly to zero and we are done. So suppose that $\|f_n\| \downarrow \alpha$, where $\alpha > 0$. If $\{F_{f_n}\}$ is not weakly null, then some subsequence $\{F_{f_{n_k}}\}$ converges weakly to $F \neq 0$ (since uniformly smooth spaces are reflexive). We shall now derive a contradiction by assuming that $F_{f_{n_k}} \xrightarrow{w} F \neq 0$. Choose $g \in \mathcal{D}$ such that $F(g) = \beta > 0$; hence $F_{f_{n_k}}(g) > \beta/2$ for all sufficiently large k. Now we use the aforementioned fact (see (4)) that uniformly smooth spaces are uniformly Fréchet differentiable. For $s > 0$, (4) yields

$$\begin{aligned} \|f_{n_k} - sg\| &= \|f_{n_k}\| - sF_{f_{n_k}}(g) + s\varepsilon\left(\frac{f_{n_k}}{\|f_{n_k}\|}, -s\frac{g}{\|f_{n_k}\|}\right) \\ &\leq \|f_{n_k}\| - \frac{s\beta}{2} + s\varepsilon\left(\frac{f_{n_k}}{\|f_{n_k}\|}, -s\frac{g}{\|f_{n_k}\|}\right) \end{aligned}$$

for large k. Since $\|f_n\| \geq \alpha > 0$, it follows that there exists $s_0 > 0$ such that

$$\|f_{n_k} - s_0 g\| \leq \|f_{n_k}\| - \frac{s_0\beta}{4}$$

for large k. As a result,

$$\|f_{n_k}\| - \|f_{n_k+1}\| \geq t\frac{s_0\beta}{4}$$

for all sufficiently large k, which contradicts the assumption that $\|f_n\| \downarrow \alpha$. Thus $\{F_{f_n}\}$ is weakly null, and hence $\{f_n\}$ is weakly null by the WN property. □

We now look at a convergence result which does not require uniform smoothness.

Let $(X_n, \|\cdot\|_n)$ be a Banach space for each $n \geq 1$. The direct sum $(\sum_{n=1}^{\infty} \oplus X_n)_{\ell_p}$ consists of all sequences $\{x_n\}_{n\geq 1}$ with $x_n \in X_n$ such that

$$\|\{x_n\}\| = \left(\sum_{n=1}^{\infty} \|x_n\|_n^p\right)^{\frac{1}{p}} < \infty.$$

Theorem 3.2 *Suppose that each X_n is Gâteaux smooth and finite-dimensional. Then the WXGA converges weakly in $(\sum_{n=1}^{\infty} \oplus X_n)_{\ell_p}$, $1 < p < \infty$.*

Proof Suppose that $\{f_n\}$ is not weakly null. There is a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \xrightarrow{w} g \neq 0$. Write $f_{n_k} = g + x_k$ where $x_k \xrightarrow{w} 0$. By passing to a further subsequence, we may assume that $\|x_k\| \to \beta \geq 0$.

Using the fact that $(\sum_{n=1}^{\infty} \oplus X_n)_{\ell_p}$ is Gâteaux smooth, choose $\phi_0 \in \mathcal{D}$, $s_0 > 0$, and $\varepsilon_0 > 0$ such that

$$\|g\|^p - \|g - s_0\phi_0\|^p = \varepsilon_0^p.$$

Then

$$\begin{aligned}\lim_{k\to\infty}\|f_{n_k}-s_0\phi_0\|^p &= \lim_{k\to\infty}\|g-s_0\phi_0+x_k\|^p\\ &= \|g-s_0\phi_0\|^p+\beta^p\\ &\le \|g\|^p-\varepsilon_0^p+\beta^p\\ &= \lim_{k\to\infty}\|f_{n_k}\|^p-\varepsilon_0^p.\end{aligned}$$

Therefore, we can choose $t_0>0$ which depends on the weakness parameter t such that

$$\|f_{n_k+1}\|^p \le \|f_{n_k}\|^p - t_0\varepsilon_0^p$$

for all sufficiently large k. However, this contradicts the fact that $\|f_n\|\downarrow\alpha>0$. □

Let us remark that characterizations of Besov spaces on [0, 1] or $\mathbb{T}$ using coefficients of wavelet or similar decompositions show that Theorem 3.2 covers the case of Besov spaces equipped with the appropriate norms.

3.3 Relaxed Greedy Algorithms

The following two greedy algorithms have been introduced and studied in [22] (see [22] for historical remarks). We begin with the Greedy Algorithm with Weakness parameter t and Relaxation $\mathbf{r}$ (GAWR($t,\mathbf{r}$)). In addition to the acronym GAWR($t,\mathbf{r}$) we will use the abbreviated acronym GAWR for the name of this algorithm. We give a general definition of the algorithm in the case of a weakness sequence τ.

3.3.1 GAWR($\tau,\mathbf{r}$)

Let $\tau:=\{t_m\}_{m=1}^\infty$, $t_m\in[0,1]$, be a weakness sequence. We define $f_0:=f$ and $G_0:=0$. Then for each $m\ge1$ we inductively define

1). $\varphi_m\in\mathcal{D}$ is any element of $\mathcal{D}$ satisfying

$$F_{f_{m-1}}(\varphi_m)\ge t_m\sup_{g\in\mathcal{D}}|F_{f_{m-1}}(g)|.$$

2). Find $\lambda_m\ge0$ such that

$$\|f-((1-r_m)G_{m-1}+\lambda_m\varphi_m)\| = \inf_{\lambda\ge0}\|f-((1-r_m)G_{m-1}+\lambda\varphi_m)\|$$

and define

$$G_m:=(1-r_m)G_{m-1}+\lambda_m\varphi_m.$$

3). Denote

$$f_m:=f-G_m.$$

In the case $\tau = \{t\}$, $t \in (0, 1]$, we write t instead of τ in the notation. We note that in the case $r_k = 0$, $k = 1, \dots,$ when there is no relaxation, the GAWR($\tau, \mathbf{0}$) coincides with the Weak Dual Greedy Algorithm [20, p. 66] (see also [6, p. 491]). We will also consider here a relaxation of the X-Greedy Algorithm (see [20, p. 39]) discussed above that corresponds to $\mathbf{r} = \mathbf{0}$ in the definition that follows.

3.3.2 X-Greedy Algorithm with Relaxation $\mathbf{r}$ *(XGAR(*$\mathbf{r}$*))*

We define $f_0 := f$ and $G_0 := 0$. Then for each $m \geq 1$ we inductively define

1). $\varphi_m \in \mathcal{D}$ and $\lambda_m \geq 0$ are such that

$$\|f - ((1 - r_m)G_{m-1} + \lambda_m \varphi_m)\| = \inf_{g \in \mathcal{D}, \lambda \geq 0} \|f - ((1 - r_m)G_{m-1} + \lambda g)\|$$

and

$$G_m := (1 - r_m)G_{m-1} + \lambda_m \varphi_m.$$

2). Denote

$$f_m := f - G_m.$$

We note that practically nothing is known about convergence and rate of convergence of the X-Greedy Algorithm. The following convergence result was proved in [22].

Theorem 3.3 *Let a sequence* $\mathbf{r} := \{r_k\}_{k=1}^{\infty}$, $r_k \in [0, 1)$, *satisfy the conditions*

$$\sum_{k=1}^{\infty} r_k = \infty, \quad r_k \to 0 \text{ as } k \to \infty.$$

Then the GAWR($t, \mathbf{r}$) *and the XGAR*($\mathbf{r}$) *converge in any uniformly smooth Banach space for each* $f \in X$ *and for all dictionaries* $\mathcal{D}$.

In this paper we discuss convergence of the GAWR($t, \mathbf{r}$) and the XGAR($\mathbf{r}$) under assumption $\sum_{k=1}^{\infty} r_k < \infty$ that is not covered by Theorem 3.3.

Proposition 3.4 *Suppose X is uniformly smooth and has the WN property. Let the relaxation* $\mathbf{r} = \{r_k\}$ *be such that* $\sum_{k=1}^{\infty} r_k < \infty$. *Then for the residual sequence* $\{f_m\}$ *of both the GAWR*($t, \mathbf{r}$) *and the XGAR*($\mathbf{r}$) *we have* $f_m \xrightarrow{w} 0$.

Proof For both algorithms we have

$$G_m = (1 - r_m)G_{m-1} + \lambda_m \varphi_m.$$

By the definition of λ_m one has

$$\|f - G_m\| \leq (1 - r_m)\|f_{m-1}\| + r_m \|f\| \tag{6}$$

and

$$\|f_{m-1}\| \geq \|f_m\| - r_m\|f\|.$$

The inequality (6) implies

$$\|f_m\| \leq \|f\|. \tag{7}$$

We now need the following simple lemma.

Lemma 3.5 *Let a sequence* $\{r_k\}$, $r_k \in [0,1]$, *be such that* $\sum_{k=1}^{\infty} r_k < \infty$. *Consider a sequence* $\{a_k\}$, $a_k \in (0, A]$, *that satisfies the inequalities*

$$a_{m-1} \geq a_m - Ar_m, \quad m = 2, 3, \dots.$$

Then either $a_m \to 0$ *as* $m \to \infty$ *or there exists an* $\alpha > 0$ *such that* $a_m \geq \alpha$ *for all* m.

Proof Suppose the sequence $\{a_m\}$ does not converge to 0. Then there exists a subsequence $\{a_{n_k}\}$ such that

$$\lim_{k\to\infty} a_{n_k} = \gamma > 0.$$

Therefore, $a_{n_k} \geq \gamma/2$ for $k \geq k_0$. For $n \in [n_{k-1}, n_k)$ one has

$$a_n \geq a_{n_k} - A\sum_{j=n+1}^{n_k} r_j. \tag{8}$$

It is clear that there exists k_1 such that for $n > n_{k_1}$ (8) implies $a_n \geq \gamma/4$. This completes the proof. □

We return to the proof of Proposition 3.4. By Lemma 3.5 we obtain from (6) and (7) that either $\|f_m\| \to 0$ or $\|f_m\| \geq \alpha > 0$. It remains to prove that the inequality $\|f_m\| \geq \alpha > 0$ implies that $\{F_{f_m}\}$ is weakly null and apply the WN property. The proof of this statement repeats the argument from the proof of Theorem 3.1. □

3.4 Modified Algorithms

First we recall the definition of the modulus of convexity $\delta_X(\varepsilon) := \delta(\varepsilon)$ of a Banach space X (see [16, pp. 59-60]):

$$\delta(\varepsilon) := \inf\left\{1 - \frac{\|x+y\|}{2} : x, y \in X, \|x\| \leq 1, \|y\| \leq 1, \|x-y\| \geq \varepsilon\right\}, \tag{9}$$

for $0 \leq \varepsilon \leq 2$. X is said to be *uniformly convex* if $\delta(\varepsilon) > 0$ for all $\varepsilon > 0$.

Suppose that X is a Banach space which has a positive modulus of convexity $\delta(s)$ for at least one $s \in (0,1)$. We consider any greedy algorithm which satisfies the following two conditions:

- the algorithm converges weakly (for every $f \neq 0$ in X, the residuals $f_n \xrightarrow{w} 0$, or in other words, $G_n \xrightarrow{w} f$);
- the norms of the residuals form a decreasing sequence $\{\|f_n\|\}$.

Now we shall modify a greedy algorithm which satisfies these two conditions. The modified algorithm generates a sequence of residuals $\{\tilde{f}_n\}$ and a corresponding sequence of approximations $\{\widetilde{G}_n\}$ such that $\|\tilde{f}_n\| \to 0$—that is, the modified algorithm converges strongly.

We proceed to describe how to apply the modified algorithm to $f = f_0 \in X$. Set $r_0 := \|f_0\|$ and $\tilde{f}_0 := f_0$. We write out the first two steps in detail and then proceed to the inductive step.

Step 1: Apply the unmodified algorithm to f_0 and generate the sequence of residuals $\{(f_0)_n\}$. We know that $(f_0 - (f_0)_n) \stackrel{w}{\to} f_0$. The norm of X is lower semicontinuous with respect to weak convergence, so

$$r_0 = \|f_0\| \le \liminf_{n\to\infty} \|f_0 - (f_0)_n\|.$$

Since $s \in (0, 1)$, we can choose n_0 such that

$$sr_0 < \|f_0 - (f_0)_{n_0}\|, \tag{10}$$

and we use n_0 to define the first residual

$$\tilde{f}_1 := \frac{f_0 + (f_0)_{n_0}}{2}. \tag{11}$$

The first approximation of f is $\widetilde{G}_1 := f_0 - \tilde{f}_1$. We will now find an upper bound for $\|\tilde{f}_1\|$ by considering the modulus of convexity $\delta(s)$. By definition, $\|\frac{f_0}{r_0}\| = 1$, and since the sequence of norms of residuals is decreasing, $\|\frac{(f_0)_{n_0}}{r_0}\| \le 1$. By (10), $\|\frac{f_0}{r_0} - \frac{(f_0)_{n_0}}{r_0}\| > s$. Therefore, by (9),

$$\delta(s) \le 1 - \left\| \frac{f_0}{2r_0} + \frac{(f_0)_{n_0}}{2r_0} \right\|,$$

which can be rewritten as

$$r_0(1 - \delta(s)) \ge \left\| \frac{f_0 + (f_0)_{n_0}}{2} \right\|. \tag{12}$$

Set $r_1 := \|\tilde{f}_1\|$. We see that $r_1 \le r_0(1 - \delta(s))$, and $0 < 1 - \delta(s) < 1$ by hypothesis.

Step 2: Apply the unmodified algorithm to $\tilde{f}_1$, generating a sequence of residuals $\{(\tilde{f}_1)_n\}$. Because $(\tilde{f}_1 - (\tilde{f}_1)_n) \stackrel{w}{\to} \tilde{f}_1$, we can choose n_1 such that

$$sr_1 < \|\tilde{f}_1 - (\tilde{f}_1)_{n_1}\|. \tag{13}$$

Set

$$\tilde{f}_2 := \frac{\tilde{f}_1 + (\tilde{f}_1)_{n_1}}{2} \quad \text{and} \quad \widetilde{G}_2 = f_0 - \tilde{f}_2.$$

We note that $\widetilde{G}_2 = (f_0 - \tilde{f}_1) + (\tilde{f}_1 - \tilde{f}_2)$.

As before, we bound $\|\tilde{f}_2\|$ using the modulus of convexity:

$$r_1(1-\delta(s)) \geq \|\tilde{f}_2\|.$$

Referring back to (12), we see that

$$r_0(1-\delta(s))^2 \geq \|\tilde{f}_2\|.$$

Set $r_2 := \|\tilde{f}_2\|$. Now we describe the inductive step.

Step k + 1: Assuming that we have found $\{\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_k\}$ and the accompanying norms $\{r_0, r_1, \ldots, r_k\}$, we apply the unmodified algorithm to $\tilde{f}_k$ in order to generate the sequence of residuals $\{(\tilde{f}_k)_n\}$. Choose n_k such that

$$sr_k < \|\tilde{f}_k - (\tilde{f}_k)_{n_k}\|.$$

Set

- $\tilde{f}_{k+1} := \frac{\tilde{f}_k + (\tilde{f}_k)_{n_k}}{2}$
- $\widetilde{G}_{k+1} := f_0 - \tilde{f}_{k+1}$
- $r_{k+1} := \|\tilde{f}_{k+1}\|$.

Finally, arguing as before,

$$r_0(1-\delta(s))^{k+1} \geq \|\tilde{f}_{k+1}\|.$$

As a result $\|\tilde{f}_n\| \to 0$ as $n \to \infty$, and the modified algorithm converges strongly.

Remark 3.6 Note that the "modification step" can and should be omitted if

$$\|(\tilde{f}_k)_{n_k}\| \leq \left\| \frac{\tilde{f}_k + (\tilde{f}_k)_{n_k}}{2} \right\|,$$

because in this case we would do better to set $\tilde{f}_{k+1} := (\tilde{f}_k)_{n_k}$ instead, and then continue with the unmodified algorithm until the next modification step. So the modification step need only be applied when it actually yields a better approximation than the unmodified algorithm. Note also that to apply the modified algorithm it is only necessary to store the value of $\tilde{f}_k$ in step $k+1$ (in addition to the updated residuals).

Remark 3.7 In all the results of this section we have applied the WN property directly, so in all results where this condition is assumed it can be dropped and the conclusion "$f_n \overset{w}{\to} 0$" replaced by "either $\|f_n\| \to 0$ or $F_{f_n} \overset{w}{\to} 0$".

4 Comments and Remarks

4.1 Banach Spaces with the WN Property

In this section we discuss which Banach spaces have the WN property. First of all it is known that $L_p[0,1]$ fails the WN property when $p \neq 2$ (see e.g. [1] for $p = 4$). For completeness we give a short proof of this fact.

Example 4.1 Fix $1 < p < \infty$ with $p \neq 2$ and let q be the Hölder conjugate index. Let $\{X_n\}$ be a sequence of independent identically distributed random variables defined on some separable probability space (Ω, Σ, P) such that

$$P(X_n = a) = 2/3 \quad \text{and} \quad P(X_n = -2a) = 1/3,$$

where $a = (3/(2 + 2^q))^{1/q}$. Then $\|X_n\|_q = 1$, and, since $\mathbb{E}[X_n] = 0$, $\{X_n\}$ is a monotone basic sequence in $L_q(\Omega)$ (by Jensen's inequality). Since $L_q(\Omega)$ is a reflexive Banach space, it follows that $\{X_n\}$ is weakly null. Now X_n is the norming functional for $Y_n \in L_p(\Omega)$, where $Y_n = |X_n|^{q-1}\operatorname{sgn}(X_n)$. But $\mathbb{E}[Y_n] = (2 - 2^{q-1})a^{q-1}/3 \neq 0$, so $\{Y_n\}$ is not weakly null in $L_p(\Omega)$. This shows that $L_p(\Omega)$ does not have the WN property.

Theorem 4.2 *Suppose X is a reflexive Banach space with the uniform Opial property. Then X has the WN property.*

Proof Assume it is not true that $x_n \xrightarrow{w} 0$—that is, suppose $\{x_n\}$ is not weakly null, where $\{x_n\}$ is as in Definition 1.1. By passing to a subsequence and relabelling, we may assume (by reflexivity of X) that $x_n \xrightarrow{w} x$, where $x \neq 0$. Now, write $x_n = x + y_n$, where $y_n \xrightarrow{w} 0$. Once more passing to a subsequence and relabelling we may assume that $\|y_n\| \to c$. If $c > 0$ then from the uniform Opial property we get

$$\lim_{n\to\infty}\left\|\frac{x_n}{c}\right\| = \lim_{n\to\infty}\left\|\frac{x}{c} + \frac{y_n}{c}\right\| \geq 1 + \tau(\|x\|/c),$$

which gives $1 \geq c + c\tau(\|x\|/c)$, so $c < 1$. (This is obviously also true if $c = 0$.) However, since $\{F_{x_n}\}$ is weakly null, we have

$$\begin{aligned} 1 &= \lim_{n\to\infty} F_{x_n}(x_n) \\ &= \lim_{n\to\infty} F_{x_n}(x) + \lim_{n\to\infty} F_{x_n}(y_n) \\ &= 0 + \lim_{n\to\infty} F_{x_n}(y_n) \\ &\leq \lim_{n\to\infty} \|y_n\| \leq c < 1, \end{aligned}$$

which is a contradiction. Therefore, $\{x_n\} \subset S(X)$ is weakly null whenever $\{F_{x_n}\}$ is weakly null, as we wished to show. □

Our next result connects bases with the uniform Opial property.

Proposition 4.3 *A Banach space with a URM basis has the uniform Opial property.*

Proof Let us take $y_n \xrightarrow{w} 0$ with $\|y_n\| = 1$ for $n = 1, 2, \ldots$ and $x \neq 0$. Let us fix a small $\delta > 0$ and N such that $\|Q_N(x)\| \leq \delta$. We have

$$\liminf_{n\to\infty}\|x + y_n\| = \liminf_{n\to\infty}\|P_N(x) + Q_N(y_n) + Q_N(x) + P_N(y_n)\|$$

$$\geq \liminf_{n\to\infty}(\|P_N(x)+Q_N(y_n)\| - \|P_N(y_n)\| - \|Q_N(x)\|)$$
$$\geq \liminf_{n\to\infty}(\|Q_N(y_n)\| + \theta(\|P_N(x)\|)) - \limsup_{n\to\infty}\|P_N(y_n)\| - \delta.$$

Since $y_n \xrightarrow{w} 0$ we have $\lim_{n\to\infty}\|P_N(y_n)\| = 0$ and $\lim_{n\to\infty}\|Q_N(y_n)\| = 1$. We also have $\|P_N(x)\| \geq \|x\| - \delta$. Putting all this together we get

$$\liminf_{n\to\infty}\|x+y_n\| \geq 1 + \theta(\|x\| - \delta).$$

Since δ was arbitrary we get the uniform Opial condition with $\tau(\varepsilon) = \theta(\varepsilon)$. □

Proposition 4.4 *Suppose X is a uniformly convex Banach space with a 1-unconditional basis $\{e_i\}$. Then the basis is URM.*

Proof Fix $0 < \varepsilon \leq 1$. To verify the URM condition, suppose that $x < y$, $\|x\| \geq \varepsilon$, and $\|y\| \leq 1$. If $\|x\| > 2$ then $\|x+y\| > 2$ by 1-unconditionality, and hence $\|y\| \leq 1 \leq \|x+y\| - 1$. So we may assume that $\|x\| \leq 2$. By 1-unconditionality,

$$\varepsilon \leq \|x\| \leq \|x+y\| = \|-x+y\|,$$

and since $y = \frac{1}{2}((x+y)+(-x+y))$ it follows that

$$\|y\| \leq \|x+y\|\left(1 - \delta\left(\frac{2\|x\|}{\|x+y\|}\right)\right)$$
$$\leq \|x+y\| - \varepsilon\delta\left(\frac{2\varepsilon}{3}\right).$$

Hence $\theta(\varepsilon) \geq \varepsilon\delta(2\varepsilon/3)$. □

There are many examples of Banach spaces which satisfy the hypotheses of Proposition 4.4. One obvious example is ℓ_p. Other examples are described in Sect. 4.1.1 below.

4.1.1 Examples of Banach Spaces with the Uniform Opial Property and URM Bases

Some examples of spaces with the uniform Opial property are presented in [14]. Now we will discuss the case of Orlicz sequence spaces. We thank Anna Kamińska for describing to us how to construct the Orlicz sequence spaces in Example 4.5. For the relevant definitions and for more general theorems which imply the facts stated below, see Chen [2] and Lindenstrauss and Tzafriri [15, Chap. 4].

Example 4.5 (Orlicz sequence spaces) Let M be an Orlicz function, M^* its complement function, and l_M its associated Orlicz sequence space. We equip l_M with the Luxemburg norm.

If M satisfies the Δ_2 condition at zero and $M(1) = 1$ then the unit vector basis $\{e_i\}$ of l_M is a normalized 1-unconditional Schauder basis for l_M. We now state some important facts which ensure that l_M is both uniformly smooth and uniformly convex [2, 13]:

- l_M is uniformly convex if and only if M and M^* satisfy Δ_2 and the function M is uniformly convex.
- l_M is uniformly smooth if and only if M and M^* satisfy Δ_2 and the function M^* is uniformly convex.
- If M and M^* satisfy Δ_2, then there exists M_1 such that M_1 and M_1^* are uniformly convex, and M_1 is equivalent to M [2, Theorem 1.18, p. 12].

Putting these facts together, we can state the following. Suppose $M(1) = 1$. The Orlicz sequence space l_M is uniformly smooth and uniformly convex if and only if M and M^* satisfy Δ_2 and are uniformly convex. In addition, $\{e_i\}$ is a normalized 1-unconditional basis. Orlicz sequence spaces with the uniform Opial property are described in [4] and more general sequence spaces in [3].

Example 4.6 $L_p[0, 1]$, $1 < p < \infty$, equipped with the square-function norm.

Let $f = \sum_{n=0}^{\infty} a_n h_n$ be the expansion of f with respect to the Haar basis. The norm defined by

$$|||f||| = \left[\int_0^1 \Big(\sum_{n=0}^{\infty} a_n^2 h_n(t)^2 \Big)^{\frac{p}{2}} dt \right]^{\frac{1}{p}}$$

is equivalent to the usual L_p norm by classical square function inequalities. Clearly, $(L_p[0, 1], |||\cdot|||)$ is isometric to a subspace of the Lebesgue-Bochner space $L_p([0, 1], \ell_2)$, which in turn is isometric to a subspace of $L_p([0, 1], L_p[0, 1]))$ since ℓ_2 is isometric to a subspace of $L_p[0, 1]$ (e.g. as the span of a sequence of independent identically distributed mean zero normal random variables). But $L_p([0, 1], L_p[0, 1])$ is naturally isometrically isomorphic to $L_p([0, 1]^2)$ (by Fubini's theorem), which in turn is isometrically isomorphic to $L_p[0, 1]$ because $[0, 1]$ and $[0, 1]^2$ have isomorphic measure algebras. In conclusion, $(L_p[0, 1], |||\cdot|||)$ is isometric to a subspace of $L_p[0, 1]$. Moreover, $(L_p[0, 1], |||\cdot|||)$ contains a subspace isometric to ℓ_p spanned by disjointly supported Haar functions. Since ℓ_p and $L_p[0, 1]$ have identical moduli of convexity and smoothness, it follows that $L_p[0, 1]$ and $(L_p[0, 1], |||\cdot|||)$ also have identical moduli of convexity and smoothness. In particular, $(L_p[0, 1], |||\cdot|||)$ is uniformly convex and uniformly smooth. Finally, $\{h_n\}_{n=0}^{\infty}$ is obviously a 1-unconditional basis for $(L_p[0, 1], |||\cdot|||)$.

4.2 Property Γ

Let us recall that this property was used by Ganichev and Kalton [8] to prove the strong convergence of the WDGA. Our aim in this section is to give an example of a Banach space with a URM basis (hence with the WN property by Theorem 4.2 and Proposition 4.3) failing property Γ. On the other hand it is proven in [8] that $L_p[0, 1]$ has property Γ but (by Example 4.1) does not have the WN property for $p \neq 2$.

Definition 4.7 A smooth Banach space X has property Γ if there is a constant $\beta > 0$ such that for any $x, y \in X$ such that $F_x(y) = 0$, we have

$$\|x + y\| \geq \|x\| + \beta F_{x+y}(y). \tag{14}$$

If we define $\varphi(t) = \|x + ty\| - \|x\|$ then we know that

$$F_{x+\alpha y}(y) = \lim_{h\to 0} \frac{\|x + (\alpha + h)y\| - \|x + \alpha y\|}{h} = \varphi'(\alpha),$$

so we can rewrite (14) as

$$\varphi(\alpha) \geq \beta\alpha\varphi'(\alpha). \tag{15}$$

The following lemma is proved for the sake of completeness. The result follows from the work of Ganichev and Kalton [9] and is included here with their permission.

Lemma 4.8 *If X is smooth and has property Γ then it is strictly convex.*

Proof Assume that X is not strictly convex. So there are points $x_0 \neq x_1$ such that $\|\lambda x_0 + (1-\lambda)x_1\| = 1$ whenever $0 \leq \lambda \leq 1$. Clearly $F_{x_0} = F_{x_1}$, so for $y = x_1 - x_0$ we have $F_{x_1}(y) = F_{x_0}(y) = 0$. We consider the function $\varphi(t) = \|x_0 + ty\| - \|x_0\|$ for $t \geq 0$. Clearly φ is increasing, differentiable, and satisfies $\varphi(t) = 0$ for $0 \leq t \leq 1$. From (15) we get

$$\varphi(t) \geq \beta t\varphi'(t) \geq \beta\varphi'(t) \tag{16}$$

for all $t \geq 1$. Suppose $\varphi(t_0) = 0$ for some $t_0 \geq 1$. Integration of (16) yields

$$\frac{\beta}{2}\varphi(t_0 + \beta/2) \geq \beta \int_{t_0}^{t_0+\beta/2} \varphi'(t)\,dt = \beta\varphi(t_0 + \beta/2),$$

and hence $\varphi(t_0 + \beta/2) = 0$. By a repeated application of the latter, starting with $t_0 = 1$, we get $\varphi(t) = 0$ for all t, which is impossible since $\varphi(t) \to \infty$ as $t \to \infty$. □

Now let is fix a concave C^1 function H on $[0, 1)$ which has the following properties:

$$H(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 0.1, \\ \sqrt{1-t^2} & \text{if } \frac{1}{\sqrt{2}} \leq t \leq 1, \end{cases}$$

and

$$H(t) < 1 \quad \text{if } t > 0.1.$$

We define a Banach space E as $\mathbb{R}^2$ equipped with the norm $\|\cdot\|$ such that

$$\|(x, y)\| \leq 1 \iff |y| \leq H(|x|).$$

The Banach space E is uniformly smooth (because H is C^1) and the basis $e_1 = (1, 0), e_2 = (0, 1)$ is a 1-unconditional URM basis. From Lemma 4.8 we see that E does not have property Γ. If we want to have an infinite-dimensional example it is enough to take $X_p = (\sum_{n=1}^{\infty} E)_p$ with $1 < p < \infty$. One easily checks that the natural basis in X_p is a 1-unconditional URM basis. Clearly X_p does not have property Γ and it is known that it is uniformly smooth (because E is, see [7]).

4.3 Dictionary Dual Greedy Step

In this section we want to discuss the following greedy step (for an otherwise unspecified greedy algorithm) which we call the *dictionary dual* step. For a dictionary $\mathcal{D} \subset X$, weakness parameter $0 < t \leq 1$, and for $x \in X$ we choose $g_0 \in \mathcal{D}$ to satisfy

$$F_{g_0}(x) \geq t \sup_{g \in \mathcal{D}} F_g(x). \tag{17}$$

This greedy step was used in [21] for a special dictionary. One could also consider the algorithm obtained by replacing (1) in the definition of the WDGA (see Sect. 3.1) by (17).

Since in the case of a Hilbert space we have

$$F_g(x) = \langle g, x \rangle = \|x\| \left\langle g, \frac{x}{\|x\|} \right\rangle = \|x\| F_x(g),$$

we see that (when $t = 1$) the dictionary dual greedy step in a Hilbert space coincides with the dual greedy step. So the dictionary dual greedy step seems to be a good generalization of the pure greedy algorithm from the Hilbert space to the Banach space setting. Indeed, one could argue that it is easier to have the functionals F_g computed once for each g belonging to the given dictionary than to have to compute F_{x_n} each time for vectors x_n which may be arbitrary (as we must do in dual greedy algorithms). The point which we want to make in this section is that for *general* dictionaries the dual greedy step (17) may create serious problems.

Definition 4.9 We call a dictionary $\mathcal{D}$ a double dictionary if $\{F_g : g \in \mathcal{D}\} \subset X^*$ is a dictionary (i.e. linearly dense) for X^*.

The Haar and Walsh systems are double dictionaries in $L_p[0,1]$ for $1 < p < \infty$. Every dictionary in a Hilbert space is a double dictionary since Hilbert spaces are self-dual.

Proposition 4.10 *Let X be separable, reflexive, smooth, strictly convex, and not isometric to a Hilbert space with* $\dim X \geq 3$. *Then X contains a countable non-double dictionary.*

Proof First note that the assumptions on X also apply to X^* by the well-known duality between smoothness and strict convexity. Let us take $V \subset X$ a subspace of codimension 1 and let $x_0^* \in X^*$ be a functional of norm 1 with $\ker x_0^* = V$. If the set $\{F_v\}_{v \in V}$ is not linearly dense in X^* then there exists a subspace $Z \subset X^*$ of codimension 1 such that $\{F_v\}_{v \in V} \subset Z$. This implies that Z norms V, i.e. for $v \in V$ we have

$$\|v\| = \sup_{z \in S_Z} |z(v)|. \tag{18}$$

Let us consider a map $q : X \to Z^*$ given by $x \mapsto x \mid Z$. From (18) we get that $q \mid V$ is an isometry from V into Z^*. Actually it must be onto Z^* because q is onto and has one-dimensional kernel. This means that $(q \mid V)^{-1} \circ q$ is a well defined norm-one operator form X onto V which is the identity on V, so it is a norm-one projection.

Now if X satisfies the assumptions of the theorem then by a theorem of James [11] X contains a one-codimensional subspace V which is not 1-complemented in X. Thus we infer that the set $\{F_v\}_{v\in V}$ is linearly dense. From this set we can choose a countable dictionary $\mathcal{D}$ which is also linearly dense by separability of X. Since $F_{F_x} = x$ (by smoothness of X^*) we see that $\{F_g : g \in \mathcal{D}\} \subset V$. This means that X^* contains a non-double dictionary. Thus, X also contains a non-double dictionary by the self-duality of our assumptions. □

Now suppose that we have a non-double dictionary $\mathcal{D}$ in a Banach space X. We have just proved in Proposition 4.10 that such a situation is quite common. Let us fix $x_0 \in X$, $x_0 \neq 0$ such that $F_g(x_0) = 0$ for all $g \in \mathcal{D}$. When we apply the dictionary dual greedy step we may choose an *arbitrary* element $g_1 \in \mathcal{D}$. We do this and put $x_1 = x_0 - \lambda g_1$ with the coefficient λ computed accordingly to our particular algorithm. Applying (17) to x_1 we get

$$|F_{g_1}(x_1)| = |\lambda| = \sup_{g\in\mathcal{D}} |F_g(x_1)|,$$

so g_1 is again an allowed (actually the best) choice. This means that using (17) we can not approximate x_0. Note that this problem appears already in nice finite-dimensional spaces (see Proposition 4.10).

To complement the previous result let us construct a two-dimensional example. Let us considers ℓ_p^2, $1 < p < \infty$, the space $\mathbb{R}^2$ with the norm

$$\|(x, y)\|_p = \big(|x|^p + |y|^p\big)^{1/p}.$$

Let us recall when we have equality in Hölder's inequality. We have

$$x_1x_2 + y_1y_2 = \big(|x_1|^p + |y_1|^p\big)^{1/p} \cdot \big(|x_2|^p + |y_2|^p\big)^{1/p}$$

if and only if $x_2 = C(\operatorname{sgn} x_1)|x_1|^{p-1}$ and $y_2 = C(\operatorname{sgn} y_1)|y_1|^{p-1}$. This means that if $a = (x, y)$ then $F_a = C(\operatorname{sgn} x \cdot |x|^{p-1}, \operatorname{sgn} y \cdot |y|^{p-1})$.

Lemma 4.11 *Let $p \neq 2$. For any vector $a = (x, y)$ such that $x \cdot y \neq 0$ and $|x| \neq |y|$ there are vectors g_1, g_2 such that*:

$$F_{g_1}(a) = 0; \tag{19}$$

$$F_a(g_2) = 0; \tag{20}$$

$$F_{g_2}(a) \neq 0. \tag{21}$$

Proof Without loss of generality $a = (1, y)$ with $y > 0$ and $y \neq 1$. Writing (19) for $g_1 = (a_1, b_1) \neq 0$ (and assuming $a_1 \geq 0$) we get

$$a_1^{p-1} + y(\operatorname{sgn} b_1)|b_1|^{p-1} = 0,$$

so a solution is given by $g_1 = (1, -y^{-1/(p-1)})$. Analogously we solve (20) for $g_2 = (a_2, b_2)$ and we get a solution $g_2 = (1, -y^{1-p})$. Using those values we write the left-hand side of (21) as

$$1 - y^{1-(p-1)^2}.$$

Since $y \neq 1$ and the exponent of y is nonzero (because $p \neq 2$) we get (21). □

Now if we apply the dictionary dual greedy step (17) in ℓ_p^2 to vector a and dictionary $\mathcal{D} = \{g_1, g_2\}$ as given in Lemma 4.11 we must choose g_2. But from (20) we get $\|a + \lambda g_2\| > \|a\|$ whenever $\lambda \neq 0$. So any algorithm producing a decreasing sequence $\{f_n\}$ (as we assumed in Sect. 3.4) and using a dictionary dual greedy step cannot converge.

Remark 4.12 We can make an infinite-dimensional example in ℓ_p by taking a dictionary $g_1, g_2, e_3, e_4, \ldots$. This is a basis and a double dictionary. Clearly, the vector $(x, y, 0, 0, \ldots)$ creates the same problem that was described above.

References

1. Browder, F.E.: Fixed point theorems for nonlinear semi-contractive mappings in Banach spaces. Arch. Ration. Mech. Anal. **21**, 259–269 (1966)
2. Chen, S.: Geometry of Orlicz spaces. With a preface by Julian Musielak. Dissertationes Math. (Rozprawy Mat.) **356** (1996)
3. Cui, Y., Hudzik, H.: Maluta's coefficient and Opial's properties in Musielak–Orlicz sequence spaces equipped with Luxemburg norm. Nonlinear Anal. **35**, 475–485 (1999)
4. Cui, Y., Hudzik, H., Yu, F.: On Opial property and Opial modulus for Orlicz sequence spaces. Nonlinear Anal. **55**, 335–350 (2003)
5. Deville, R., Godefroy, G., Zizler, V.: Smoothness and Renormings in Banach Spaces. Pitman Monographs in Pure and Applied Mathematics, vol. 64. Longman Scientific & Technical, Harlow. Copublished in the United States by Wiley, New York (1993)
6. Dilworth, S.J., Kutzarova, D., Temlyakov, V.N.: Convergence of some greedy algorithms in Banach spaces. J. Fourier Anal. Appl. **8**, 489–505 (2002)
7. Figiel, T.: On the moduli of convexity and smoothness. Studia Math. **56**, 121–155 (1976)
8. Ganichev, M., Kalton, N.J.: Convergence of the weak dual greedy algorithm in L_p-spaces. J. Approx. Theory **124**, 89–95 (2003)
9. Ganichev, M., Kalton, N.J.: Convergence of the dual greedy algorithm in Banach spaces. Preprint (2007)
10. Huber, P.J.: Projection pursuit. Ann. Stat. **13**, 435–475 (1985)
11. James, R.C.: Inner products in normed linear spaces. Bull. Am. Math. Soc. **53**, 559–566 (1947)
12. Jones, L.: On a conjecture of Huber concerning the convergence of projection pursuit regression. Ann. Stat. **15**, 880–882 (1987)
13. Kamińska, A.: Personal communication, April 6, 2006
14. Khamsi, M.A.: On uniform Opial condition and uniform Kadec-Klee property in Banach and metric spaces. Nonlinear Anal. **26**, 1733–1748 (1996)
15. Lindenstrauss, J., Tzafriri, L.: Classical Banach Spaces I, Sequence Spaces. Ergebnisse der Mathematik, vol. 92. Springer, Berlin (1977)
16. Lindenstrauss, J., Tzafriri, L.: Classical Banach spaces II, Function Spaces. Ergebnisse der Mathematik, vol. 97. Springer, Berlin (1979)
17. Prus, S.: Banach spaces with the uniform Opial property. Nonlinear Anal. **18**, 697–704 (1992)

18. Prus, S.: Geometrical background of metric fixed point theory. In: Kirk, W.A., Sims, B. (eds.) Handbook of Metric Fixed Point Theory, pp. 93–132. Kluwer Academic, Dordrecht (2001)
19. Temlyakov, V.N.: Greedy algorithms in Banach spaces. Adv. Comput. Math. **14**, 277–292 (2001)
20. Temlyakov, V.N.: Nonlinear methods of approximation. Found. Comput. Math. **3**, 33–107 (2003)
21. Temlyakov, V.N.: Greedy approximations. In: Found. Comput. Math., Santander 2005. London Math. Soc. Lecture Note Ser., vol. 331, pp. 371–394. Cambridge University Press, Cambridge (2006)
22. Temlyakov, V.N.: Relaxation in greedy approximation. Constr. Approx. **28**, 1–25 (2008)